Control of Cognitive Processes

Attention and Performance XVIII

edited by Stephen Monsell and Jon Driver

This book is based on the papers presented at the Eighteenth International Symposium on Attention and Performance, held at Cumberland Lodge, The Great Park, Windsor, Berkshire, England, July 12–18, 1998.

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

This book was set in Palatino on 'Quark' by Asco Typesetters, Hong Kong and was printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

International Symposium on Attention and Performance (XVIII : 1998 : Berkshire, England)
Control of cognitive processes : attention and performance XVIII / edited by Stephen Monsell and Jon Driver.
p. cm
"Based on the papers presented at the Eighteenth International Symposium on Attention and Performance, held at Cumberland Lodge, The Great Park, Windsor, Berkshire, England, July 12–18, 1998."—CIP T.p.
ISBN 0-262-13367-9 (hc : alk. paper)
1. Cognition—Congresses. I. Monsell, Stephen. II. Driver, Jon. III. Title.
BF311.157 1998
153.7'33—dc21 00-030732

4 On the Time Course of Top-Down and Bottom-Up Control of Visual Attention

Jan Theeuwes, Paul Atchley, and Arthur F. Kramer

ABSTRACT Previous research showed that a salient feature singleton captured attention bottom-up (Theeuwes 1991a, 1992, 1994a). A salient color singleton interfered with search for a less salient shape singleton, which suggested that early processing was driven by bottom-up saliency factors. The present experiments examined how bottom-up and top-down processing develops over time. Subjects searched for a shape singleton target and had to ignore a color singleton distractor presented at different stimulus onset asynchronies prior to the search display. The results indicate that when the target and distractor were presented simultaneously, the salient singleton distractor captured attention, whereas when the distractor singleton was presented about 150 msec before the target singleton, the distractor did not disrupt performance. The findings suggest a stimulus-driven model of selection in which early processing is solely driven by bottom-up saliency factors. In later processing, the early bottom-up activation of the distractor can be overridden by top-down attentional control.

One of the most basic questions in the study of attention is the extent to which top-down attentional control can prevent distraction from irrelevant stimuli. Visual selective attention is thought to be (1) *goal directed* when attentional priority is given to only those objects and events that are in line with the current goals of the observer: and (2) *stimulus driven* when, irrespective of the intentions or goals of the observer, objects and events involuntarily receive attentional priority—a phenomenon referred to as "attentional capture" (for recent reviews, see Egeth and Yantis 1997; Theeuwes 1993, 1994b; Yantis, 1993, 1996). These two mechanisms of selection have been referred to as "top-down" and "bottom-up" attentional control, respectively, (e.g., Eriksen and Hoffman 1972; Posner 1980; Theeuwes 1991b; Yantis and Jonides 1984).

Many models of visual search assume that visual selection is the result of an interaction between goal-directed and stimulus-driven factors (e.g., Cave and Wolfe 1990; Treisman and Sato 1990). Typically, it is assumed that bottom-up activation occurs during early preattentive processing in which the visual field is segmented into functional perceptual units. Bottom-up activation is a measure of how salient an item is in its context. An item that is locally unique in some basic visual dimension—usually referred to as *a* "feature singleton" or simply a "singleton"—will generate a large bottom-up activation (e.g., a red poppy in a field of green grass). Top-down activation may also operate during attentional processing. Various studies have demonstrated that in more complex search tasks, knowledge of the specific task demands may guide attention to only those locations that match the target-relevant feature. For example, Kaptein, Theeuwes, and van der Heijden (1995) showed that when searching for a red vertical line segment between red tilted and green vertical line segments, subjects searched serially among the red items while ignoring the green line segments (see also Egeth, Virzi, and Garbart 1984). Top-down guidance is typically assumed to proceed either by *activation* of features that match those of the target (e.g., Wolfe 1994) or by *inhibition* of features that do not (Treisman and Sato 1990).

In a series of experiments, Theeuwes (1991a, 1992, 1996) showed that a salient feature singleton captured attention bottom-up. Even though subjects had a clear top-down attentional set to search for a particular singleton, performance was disrupted by a distractor with a salient, unique feature in a task-irrelevant dimension. Top-down control of attention could not entirely override bottom-up interference from a singleton distractor known to be irrelevant. For example, Theeuwes (1992) presented subjects with displays consisting of colored circles or diamonds appearing on the circumference of an imaginary circle. Line segments of different orientations appeared in the circles and diamonds. Subjects had to determine the orientation of the line segment appearing in the target shape. Subjects searched for a shape singleton, a single green diamond among green circles. Time to find the shape singleton increased when an irrelevant color singleton was present (i.e., one of the circles was red). Even though subjects had a clear top-down set to search for the shape singleton (i.e., the single green diamond), the presence of an irrelevant singleton (i.e., the single red circle) caused interference. It was shown that selectivity depended on the relative salience of the stimulus attributes: when the color singleton was made less salient than the shape singleton (by reducing the color difference between the target and the nontarget elements), the shape singleton interfered with search for the color singleton, whereas the color singleton no longer interfered with the search for the shape singleton.

Based on these experiments, Theeuwes (1991a, 1992, 1994a, 1996) concluded that early preattentive processing is driven by bottom-up factors such as salience. Attention is captured by the most salient singleton in the display, regardless of whether the property defining that singleton is relevant for the task or not (for more recent evidence, see Bacon and Egeth 1994, exp. 1; Caputo and Guerra 1998; Joseph and Optican 1996; Kawahara and Toshima 1996; Kim and Cave 1999; Kumada 1999; Todd and Kramer 1994). When engaged in parallel search for a particular feature singleton (e.g., a diamond among circles), the extent to which singletons capture attention is determined by the relative salience of the singletons present in the visual field. It was suggested that, irrespective of any top-down control, spatial attention is automatically and involuntarily captured by the most salient singleton. The shift of spatial attention to the location of the singleton implies that the singleton is selected for further processing. If this singleton is the target, a response is made. If it is not the target, attention is directed to the next most salient singleton. The initial shift of attention to the most salient singleton is thought to be the result of relatively inflexible, "hard-wired" mechanisms, triggered by the presence of these difference signal interrupts. Consistent with proposals by Sagi and Julesz (1985) and Koch and Ullman (1985), it is assumed that the parallel process can only perform *local mismatch* detection (i.e., indicating the presence of a discontinuity, but not its nature) followed by a serial stage directed to areas of the visual field with the largest magnitude mismatches.

Contrary to these findings, a group of other researchers have claimed that the ability of a singleton to capture attention is contingent on whether an attention-capturing stimulus is consistent with top-down settings established "off-line" on the basis of current attentional goals (Folk, Remington, and Johnston 1992; Folk and Remington 1998). According to this "contingent capture" model, only stimuli that match the top-down control settings will capture attention; stimuli that do not match the top-down settings will be ignored. Top-down control is thus possible even when target and distractor are both singletons. Along these lines, it was argued that in Theeuwes's experiments the irrelevant singleton captured attention because subjects were set to find a singleton (e.g., a local mismatch) rather than a particular feature, such as a red circle (see Bacon and Egeth 1994). It was claimed that irrespective of the bottom-up saliency, the singleton that matched the top-down setting would capture attention. These claims are based on evidence from experiments in which subjects had to ignore a cue that appeared 150 msec before the presentation of the target display (Folk, Remington, and Johnston 1992). Subjects responded to a character shape (*X* versus =) that, in different conditions, had either a unique color or a unique abrupt onset. When the search display was preceded by a to be ignored featural singleton (the cue) that matched the singleton for which they were searching, the cue captured attention as evidenced by a prolonged reaction time to identify the target (i.e., when the cue and target appeared in different spatial locations). On the other hand, if the to be ignored featural singleton cue did not match the singleton for which they were searching, its appearance apparently did not capture attention. This "contingent" capture of attention occurred for both color and onset conditions, and is considered evidence that involuntary capture is contingent on the adoption of some attentional set.

The critical finding in these studies is that a cue that does not match the top-down search goal (i.e., the defining property of the target) does not affect response time (RT), whereas a cue that matches the search goal does. In other words, if subjects were searching for a red plus sign, they were more likely to be distracted by a red cue than by an abrupt onset cue, and vice versa. Folk, Remington, and Johnston (1992) have suggested that the absence of an effect on RT for a cue that does not match the target indicates that the cue did not capture attention. On the other hand, the irrelevant cue may indeed have captured attention, but because the cue display came on 150 msec before the search display, subjects may have been able to overcome the attentional capture by the time the search display was presented (see also Theeuwes 1994a,b). Disengagement of attention from the cue may have been relatively fast when the cue and target did not share the same defining properties (e.g., the cue is red and the target is an onset), whereas disengagement from the cue may have been relatively slow in the case where the cue and target share the same defining properties (e.g., both were red). Such a mechanism could explain why there are RT costs when the cue and target have the same defining characteristics—and no costs when cue and target are different. This does not imply, however, that there is no capture of attention by the irrelevant cue singleton; it simply indicates that, after a certain time, subjects are able to exert top-down control over the erroneous capture of attention by the irrelevant singleton, to overcome its effects.

This account holds that early preattentive processing is driven by solely bottom-up feature salience factors, generating an activation pattern on which later attentive processing may then exert control to give priority to elements that match the top-down attentional set. It thus remains consistent with the claim of Theeuwes (1991a, 1992, 1994a, 1996) that during early *preattentive* processing, top-down control is not possible. It is also in line with models of visual search suggesting that during *attentive* processing either top-down inhibition may be applied to features that match the distractors (Treisman and Sato 1990) or top-down activation, to features that match the target (Wolfe 1994).

If the model presented above is correct, it should be possible to reveal how bottom-up and top-down processing develop over time. As in previous studies (e.g., Theeuwes 1992), subjects searched multielement displays for a shape singleton and reported the letter located inside the shape singleton. On some trials, an irrelevant salient color singleton was presented along with a premask display at different stimulus onset asynchronies (SOAs) before the onset of the search display. When the target and distractor singleton are presented close in time, and attention is captured by the distractor, search for the target singleton should be slowed. If, however, the singleton distractor is presented well in advance of the search display, subjects may be able to exert top-down control over the irrelevant singleton, ensuring that, by the time of the arrival of the search display, attention is directed to the target singleton. In these latter conditions, there should be no effect of the singleton distractor on search time.

4.1 EXPERIMENT 1

A visual search task similar to that in Theeuwes 1992 was employed, where subjects had to search for a feature singleton, and where this singleton is typically detected by means of preattentive parallel search. Subjects searched for a shape singleton (a single gray diamond among eight gray circles) and had to determine the orientation of the letter *C* (*C* or reversed *C*) appearing in the diamond. Determining the orientation of the letter *C* requires the allocation of focal attention to the location of the shape singleton. In the distractor condition, one of the circles was red. Because previous studies (see Theeuwes 1991a, 1992) have shown that such a color singleton is more salient than a shape singleton, it was expected that, in line with previous studies, the color singleton (i.e., the distractor) would interfere with the search for the shape singleton (i.e., the target). To determine the time course of bottom-up and top-down activation, the singleton distractor (the red circle) appeared at different SOAs prior to the presentation of the target display.

Subjects

Twelve subjects, ranging in age from 18 to 30, participated as paid volunteers. All had self-reported normal or corrected-to-normal vision and reported having no color vision defects.

Apparatus

A 486 computer with an SVGA color monitor controlled the timing of the events, generated stimuli and recorded reaction times. The "/"-key and the "z"-key of the computer keyboard were used as response buttons. All subjects were tested in a sound-attenuated, dimly lit room, with their head resting on a chinrest. The monitor was located at eye level, 60 cm from the chinrest.

Stimuli

Subjects performed a visual search task in which they searched for a uniquely shaped element (a diamond located between circles) and responded to the letter located inside this uniquely shaped singleton. The display consisted of nine elements equally spaced around the fixation point on an imaginary circle whose radius was 3.4 degrees. In the control condition each display contained one gray outline diamond (1.4 degrees on a side) surrounded by eight gray outline circles (1.4 degrees in diameter). In the distractor condition, one of the gray outline circles was replaced by a red circle producing a condition identical to that of Theeuwes 1992, in which the target had a unique shape (shape singleton)

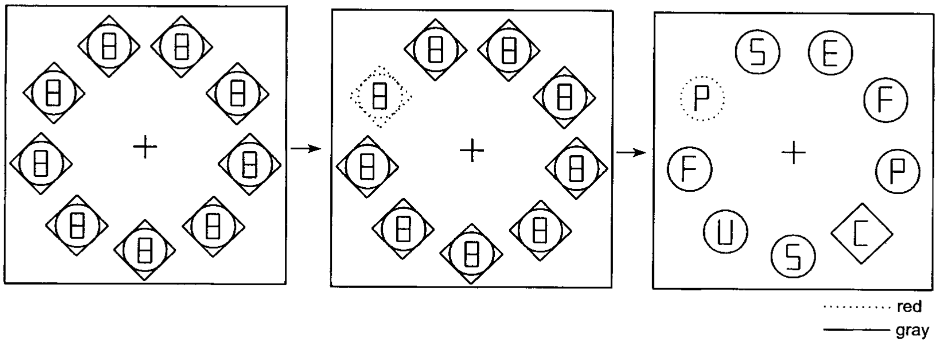

Figure 4.1 Trial events in experiment 1. The premask display (left panel) was presented for 700 msec. At stimulus onset asynchronies of 50, 100, 150, 200, 250, or 300 msec before the presentation of the search display, the color of one of the elements of the premask changed from gray into equiluminant red (middle panel). The search display (right side) contained both the color singleton distractor (the red element) and a shape target singleton (the diamond).

while the distractor had a unique color (color singleton). To ensure that distractor effects were not due to attention encompassing both the target singleton and the neighboring color singleton, the color singleton distractor was never placed adjacent to the target (i.e, there was always one gray element between the target and color distractor).

Each display element contained a letter (0.4×0.8 degrees). The uniquely shaped outline diamond (i.e., the target) contained either a *C* or a reversed *C*, the orientation of the letter determining the response (subjects pressed the "z"-key for a *C* and the "/"-key for a reversed *C*). The letters inside the other eight circles were randomly sampled from the set *E*, *P*, *F*, *U* and *S*. The letters were presented in white (11.0 cd/m^2) and the circle and diamond were presented in gray (6.4 cd/m^2). The color singleton distractor was presented in red (6.3 cd/m^2).

Design and Procedure

The sequence of events was as follows: Initially, a fixation dot was presented for 1,000 msec. Then the premask display came on consisting of nine premask elements, each composed of a single outline circle and diamond, and each containing a figure-eight premask letter (see figure 4.1). The premask display was presented for 700 msec. The color of one of the elements of the premasks changed from gray to equiluminant red with SOAs of 50, 100, 150, 200, 250, or 300 msec before the presentation of the search display. The search display was revealed by removing particular diamonds or circles of the premask display resulting in a search display consisting of eight circles and one diamond. Simultaneously with the removal of the premask, the letters inside the outline elements were

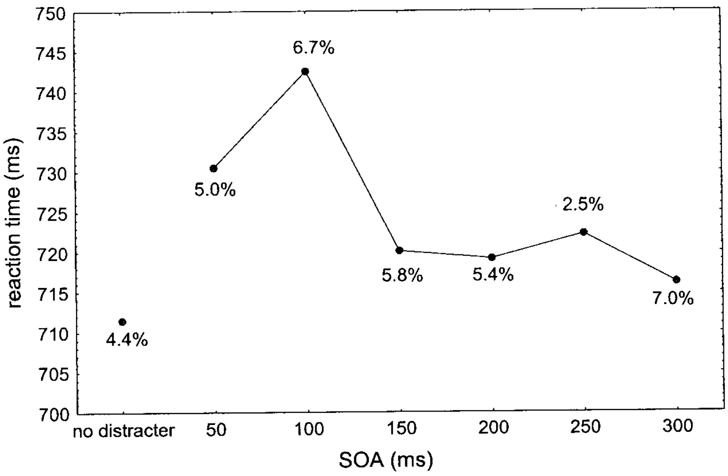

Figure 4.2 Experiment 1: Mean RTs and error percentages as a function of stimulus onset asynchrony (SOA) for the distractor and no-distractor conditions.

displayed by removing line elements from the figure eights. The search display remained present for a maximum of 2 sec until a response was emitted.

Each subject performed 240 trials, 120 no-distractor and 120 distractor trials which were presented randomly within blocks of trials. SOA between premask and search display was varied randomly between trials as well. Subjects were told to keep their eyes fixated at the fixation dot. Subjects received 240 practice trials prior to the experimental trials, as well as feedback about their performance in terms of RT and error rates after each block of 60 trials. Prior to the start of the experiment, subjects were instructed to search for the diamond and respond to the orientation of the letter inside the diamond by pressing the appropriate response key. They were told to ignore the uniquely colored red singleton.

Results

Response times longer than 1,200 msec were counted as errors, which led to a loss of less than 1% of the trials. A one-way analysis of variance (ANOVA) with no distractor or SOAs of 50, 100, 150, 200, 250, and 300 msec as levels showed a significant main effect: $F(6, 66) = 2.3$; $p < 0.05$. Additional planned comparisons showed that the RT at SOAs of 50 msec (731 msec) and 100 msec (742 msec) were significantly slower ($p < 0.05$) than the RT in the no-distractor control condition (711 msec), indicating that at the early SOA the singleton distractor interfered with search for the target singleton. However, the RTs of the later SOAs (150, 200, 250, and 300 msec) were not significantly different from the no-distractor condition suggesting that in these conditions search for the target singleton was not affected by the presence of the singleton distractor (see figure 4.2). Note that a distractor, when presented close in time to the target, slows down search by about 25 msec, an effect size very similar to that reported in Theeuwes 1992. The error rates were low (about 4.9%) and did not vary systematically with any of the conditions.

Discussion

The present results confirm earlier findings (e.g., Theeuwes 1991a, 1992) that the presence of a irrelevant salient distractor interferes with search for a relevant target singleton. The analysis of SOA suggests that there is a reliable effect of the distractor at the early SOAs (50 and 100 msec) but not at the later SOAs (150, 200, 250, 300 msec).

The results regarding SOA are in line with our predictions: at the early SOAs when distractor and target are presented in close succession, there is a clear interference effect of the distractor. It was argued that in these conditions, when target and distractor were presented in close temporal proximity, there was not enough time to exert top-down control that could have overcome attentional capture by the salient distractor. When, however, the singleton distractor was presented a considerable time (SOAs of 150 to 300 msec) before the presentation of the target singleton, sufficient top-down control could be exerted that there was no sign of attentional capture by the distractor. Indeed, response times at SOAs of 150 to 300 msec did not differ significantly from that in the no-distractor condition.

4.2 EXPERIMENT 2

Experiment 1 suggests that early in processing attention is captured by the salient distractor and that, later, attentional capture is overcome by top-down attentional control. To determine whether spatial attention was indeed captured by the distractor, we used the response congruency paradigm (Eriksen and Eriksen 1974; Eriksen and Hoffman 1972), in which subjects have to ignore a stimulus that is either congruent or incongruent with the response to the target. In previous studies (Theeuwes 1996; Theeuwes and Burger 1998; Theeuwes et al. 1999) investigating whether subjects could intentionally ignore salient but irrelevant singleton elements, the element to ignore was either identical to or different from the target element they were looking for. The results showed that the identity of the element to be ignored had an effect on response time suggesting that indeed spatial attention was directed at the location of the distractor element. Subjects were faster when the distractor element was identical to the target (congruent with the response) than when the distractor element was different from the target (incongruent with the response).

To determine whether spatial attention was shifted to the location of the color singleton distractor, we also presented a *C* or reversed *C* inside the color singleton distractor at the various SOAs used here. This letter was either identical with the letter inside the target shape singleton (and therefore congruent with the response) or different from the letter inside the target shape singleton (and therefore incongruent with the response). If attention is indeed captured by the color singleton distractor, then the identity of the letter inside the colored singleton distractor should have an effect on responding, that is, a letter congruent with the response should produce faster RTs than a letter incongruent with the response. If attention is not captured by the colored singleton, then there should be no congruency effect on RT.

Subjects

Fifteen subjects, ranging in age from 18 to 30, participated as paid volunteers.

Stimuli

The stimuli were identical to those in experiment 1. The letter located inside the irrelevant color singleton distractor was either a *C* or a reversed *C*, and this could be congruent or incongruent with the target letter inside the relevant shape singleton.

Design and Procedure

Only SOAs of 50, 100, 200, and 400 msec were used. In the current experiment, there was always a red singleton distractor present in each display. A congruent or incongruent letter was presented inside the distractor singleton.[1] Note that the letter inside the singleton distractor was revealed simultaneously with the red singleton distractor element. In other words, the letter (which could be congruent or incongruent with the response) was presented simultaneously with the singleton distractor and therefore this letter was presented 50, 100, 200, or 400 msec before the presentation of the other letters of the search display (including the target letter). SOA was varied randomly within blocks of trials. Subjects received 240 practice trials and 240 experimental trials.

Results

Response times longer than 1,300 msec were counted as errors, which led to a loss of 0.9% of the trials. An ANOVA with SOA (50, 100, 200, 400 msec) and congruency (congruent versus incongruent) as orthogonal within subject factors showed an effect of SOA: $F(3, 42) = 7.8$; $p < 0.001$;

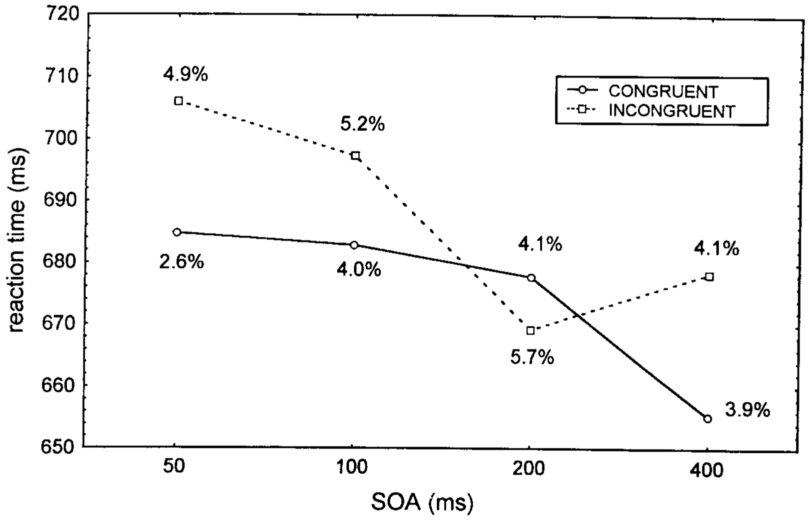

Figure 4.3 Experiment 2: Mean RTs and error percentages as a function of stimulus onset asynchrony (SOA) for the congruent and incongruent conditions.

and of congruency: $F(1, 14) = 10.3$; $p < 0.001$. The interaction between SOA and congruency was also reliable: $F(3, 42) = 3.37$; $p < 0.05$. As is clear from figure 4.3, response times become faster with increasing SOA, suggesting that (in line with experiment 1) the effect of the singleton distractor diminishes with increasing SOA. Additional planned comparisons show that a reliable congruency effect at SOA, 50 and 400 msec ($p < 0.05$) and a marginally significant congruency effect at SOA 100 msec ($p = 0.07$). At SOA 200 msec, congruency failed to reach significance ($p = 0.29$; 678 msec versus 670 msec). Also, as is clear from figure 4.3, when the letter inside the singleton distractor was congruent with the response to the letter inside the target singleton response times were faster than when it was incongruent. The finding that the letter inside the singleton distractor did affect responding to the target singleton can only be explained by assuming that at some point attention resided at the location of the singleton distractor (but see Folk and Remington 1998). The error rates were low (4.4%) and did not vary systematically with any of the conditions.

Discussion

Experiment 2 shows that response latencies become shorter with increasing SOA, suggesting again that presenting the distractor in advance of the target overcomes attentional capture by the distractor. As in experiment 1, the distractor seems to slow search by about 25 msec at the two short SOAs (50 and 100 msec).

The overall congruency effect indicates that the identity of the letter inside the singleton distractor had an effect on the response to the letter appearing inside the target singleton. When the letter inside the distractor was identical to the letter inside the target singleton, and therefore congruent with the response, response times were faster than when the letter inside the distractor was incongruent with the response to the letter inside the target singleton, a result identical to that in Theeuwes 1996. These findings are consistent with attention being *captured* by the irrelevant singleton. Because capturing attention implies that focal attention was directed to the irrelevant singleton, the identity of the letter became available, thereby affecting the speed of responding to the target.

Folk and Remington (1998) have suggested an alternative explanation for such findings. Instead of assuming that attention was captured by the irrelevant singleton, they suggested that the congruency effect as observed in Theeuwes 1996 and in Theeuwes and Burger 1998 was the result of processing the target and distractor letter in parallel. Such an explanation, though possible, is unlikely: at the eccentricities used in the current experiments, letters cannot be processed efficiently in parallel for discriminations such as C versus reversed C (see, for example, Theeuwes 1991c; Wolfe 1994). Usually, when subjects search for a target letter among nontarget letters, search time increases linearly with the number of nontarget letters in the display, a result typically seen as evidence for spatially serial search. Given these considerations, the most likely explanation is that the identity of the letter in the irrelevant singleton affected responding because attention was directed at the location of the singleton distractor before a response was made. In addition, the control experiment (see note 1), in which a congruent or incongruent letter was placed in a nonsingleton item, showed no effect of congruency ($F = 1$), providing evidence that the congruency effect only shows when attention is attracted to the location of the colored singleton. This finding suggests that parallel processing of *all* letters (including the congruent or incongruent letter placed in the nonsingleton) is highly unlikely.

It is important to note that there is a clear congruency effect at SOA 400 msec ($p = 0.0065$). This finding is important because it implies that even when the singleton distractor (with the congruent or incongruent letter inside) is presented 400 msec before the presentation of the search display, attention was captured by the singleton. In other words (as demonstrated in experiment 1), at SOAs of 200 msec, subjects had enough time between the presentation of distractor and target, not to prevent attentional capture, but to gain attentional control after their attention had been erroneously captured by the salient distractor.

Another interesting finding is that at SOA 200 msec, the congruency effect is absent (if anything, the effect is reversed). This suggests that to gain attentional control subjects may have inhibited the singleton distractor location and thereby reduced the influence of the letter inside the singleton distractor. Because of this inhibition, the letter inside the distractor no longer influences responding to the target letter. The fact that the congruency effect is absent at SOA 200 msec but not at SOA 400 msec suggests that the inhibition may be transient.

4.3 EXPERIMENT 3

The goal of experiment 3 was to investigate the possible role of inhibition of the distractor color over trials. Experiment 3 was identical to experiment 2 except that the color of the singleton distractor could be either red or green and changed randomly from trial to trial. The results of experiment 2 suggesting inhibition of the distractor at SOA 200 msec and not at SOA 400 msec implies that the inhibition may be relatively short-lived. If inhibition is relatively brief, then changing the color of the distractor from trial to trial should produce the same pattern of effects as that observed in experiment 2.

If, however, attentional set (e.g., in the sense of inhibiting a specific color) is carried over from one trial to the next, then response latencies should be faster when the singleton distractor has the same color as on the previous trial than when it does not. Such a result would be consistent with Maljkovic and Nakayama 1994, which showed that visual search responses were faster when the color of the target singleton was repeated from the previous trial than when it was changed. Subjects were considered to be relatively fast on same-color trials because they could retrieve an attentional set identical to the one used in the previous trial. Although it is not clear whether such a repetition effect also occurs when distractor rather than target colors are changed, if repeating the same attentional set produced a general effect, then a repetition effect should also be observed for the distractors in the present studies. Note that, as in Maljkovic and Nakayama 1994, any repetition effect in the current experiment cannot be a response-based effect because subjects did not respond to the color, but to the letters inside the elements.

Subjects

Twelve subjects, ranging in age from 18 to 30, participated as paid volunteers.

Stimuli

The stimuli were identical to those in experiment 2. The singleton distractor was either red or green and changed color randomly from trial to trial.

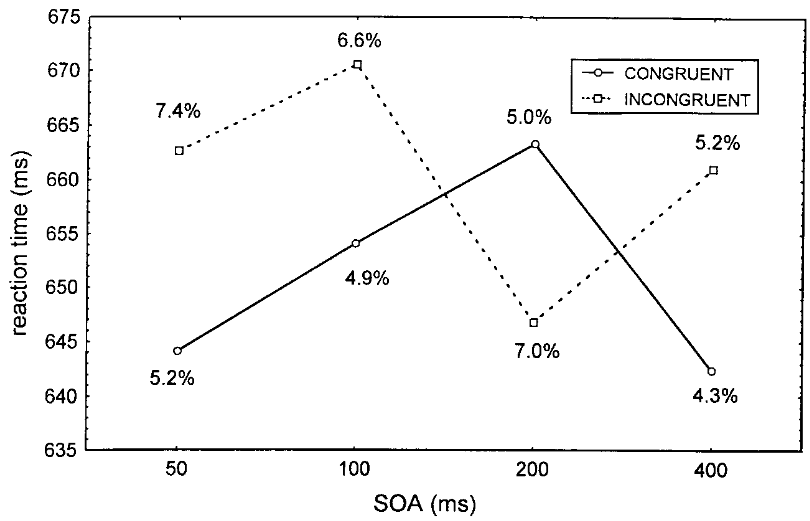

Figure 4.4 Experiment 3: Mean RTs and error percentages as a function of SOA for the congruent and incongruent conditions.

Design and Procedure

Subjects received 512 practice and 512 experimental trials.

Results

Response times longer than 1,300 msec were counted as errors, which led to a loss of 1.1% of the trials. An ANOVA with SOA (50, 100, 200, 400 msec) and congruency (congruent versus incongruent) as factors showed an effect of congruency: $F(1, 11) = 4.90$; $p < 0.05$; and of congruency × SOA: $F(3, 33) = 6.54$; $p < 0.001$. Planned comparisons indicate that, for all SOAs, the difference between the congruent and incongruent conditions is reliable (all $p < 0.05$). Note, however, that at SOA 200 msec this effect is reversed ($p = 0.02$), that is, incongruent responses are *faster* than congruent responses (see figure 4.4).

An additional analysis was carried out to determine whether changing the color of the singleton distractor over trials had an effect on response latencies. An ANOVA showed no effect of color change: $F(1, 11) = 3.0$; $p = 0.11$; nor did distractor color change interact with any of the other variables—color change × congruency: $F(1, 11) = 0.07$; color change × SOA: $F(3, 33) = 0.86$. This suggests subjects were not able to carry over the attentional set (including the color to inhibit) from the previous trial in order to speed up responding. The error rates were low (5.5%) and did not vary systematically with any of the variables.

Discussion

The finding that RTs in trials in which the distractor color switched were the same as when the color remained the same suggests that attentional set in the sense of which color to inhibit does not carry over from one trial to the next. Unlike the findings in Maljkovic and Nakayama 1994, which showed a repetition effect for the *target* color, the current findings indicate that this does not hold for the *distractor* color. The results suggest that the color to inhibit may not be part of the attentional set that transfers from one trial to the next. Note, however, that in experiment 3 the target remained fixed over trials. If specifying the target is the most important feature of the attentional set, then one may argue that repetition effects of the distractor color were not observed in experiment 3 because the target remained the same. Future studies may address whether switching the color of the distractor produces a repetition effect when the color of the target also changes from trial to trial. Overall, consistent with the findings of experiment 2 that overcoming of the distractor effect was relatively short-lived, the current findings suggest that rejection of the relevant color singleton does not transfer from one trial to the next.

The congruency effects are similar to those of experiment 2. For SOAs 50, 100, and 400 msec, there is a clear congruency effect in the sense that congruent responses are faster than incongruent responses. Yet, consistent with a trend in experiment 2, at an SOA of 200 msec, the congruency effect is reversed, that is, congruent responses are faster than incongruent. The findings suggest that in order to redirect attention away from the singleton distractor location, subjects may have inhibited the location of the distractor, and thereby inhibited the letter inside the singleton distractor. When the inhibited letter is identical to the target letter (i.e., the congruent condition), subjects are relatively slow. On the other hand, when the inhibited letter is different from the target letter (the incongruent condition), the letter is not inhibited and subjects are relatively fast.

Distractor inhibition also appears in many experiments demonstrating negative priming, in which the response to a stimulus is slowed when the previously inhibited stimulus becomes relevant for responding (e.g., Neill and Valdes 1996). For example, Tipper and Cranston (1985) showed that when subjects ignored a letter on trial n, the response to a letter with the same identity on trial $n + 1$ was impaired, a condition comparable to the congruency manipulation in experiments 2 and 3. It is hypothesized that actively inhibiting the potentially competing response from the letter in the singleton to be ignored, may cause a reversal of the congruency effect; that is, a response congruent with the letter inside the distractor is slower than a response that is incongruent. Note that this reversal only occurs when the distractor is presented 200 msec before the presentation of the target, suggesting it takes time for inhibition to accrue. A similar pattern of facilitation and inhibition appears in experiments addressing "inhibition of return": targets appearing on the cued side show an RT advantage for the first 150 msec, which is replaced by an inhibition after 250 to 300 msec (Posner and Cohen 1984).

The observation in experiment 3 that the congruency effect returns after a time interval of 400 msec is not in line with findings from either the "negative priming" nor the "inhibition of return" literature. At the early SOAs, in which distractor and target are presented within 100 msec, active top-down inhibition at the location of the distractor may start to build up, yet, before it is complete, the appearance of the target singleton causes attention to be automatically captured by the location of the target singleton. In other words, there may not be enough time to allow active top-down inhibition at the early SOA, resulting in a "typical" congruency effect, as observed in previous studies, where target and singleton distractor were presented simultaneously (see Theeuwes 1996).

At the later SOA of 200 msec, as evidenced by the absence of an interference effect of the distractor, subjects may have enough time to exert top-down control. Top-down control results in active inhibition of the singleton distractor, including the letter located inside the distractor. Inhibition is important because the distractor and target are presented in relatively close succession (i.e., within a 200 msec time frame), and will compete for attention. At this point, it is not clear why the congruency effect returns at SOA 400 msec. Perhaps it is impossible to maintain this type of inhibition over a longer time period.

4.4 GENERAL DISCUSSION

The current experiments were designed to examine the time course of bottom-up and top-down processing in visual search. The results indicate that a salient singleton distractor presented close in time to the target singleton causes interference, as demonstrated by response times that are significantly longer than those in the no-distractor condition. The finding that the letter inside the singleton distractor had an effect on responding to the target (i.e., the congruency effect) also suggests that spatial attention was drawn to the location of the distractor providing evidence that the increase in RT is indeed due to attentional capture.

When a singleton distractor is presented at least 150 msec in advance of the target, the interference effect is no longer observed, although the finding that the letter inside the singleton distractor has an effect on RT at still longer SOAs indicates that attention was captured by the singleton distractor. Yet, with an interval of 150–200 msec between the presentation of distractor and target, there was sufficient time to reorient spatial attention from the location of the distractor. When, at that point, the target singleton is presented, attention is immediately directed to the target singleton resulting in response times equivalent to those in the no-distractor condition.

When a singleton distractor is presented 150–200 msec in advance of the target, it is assumed that top-down control can reduce or eliminate the effect of the distractor. Note, however, that the presence of a congruency effect at the longer SOAs indicates that top-down attentional control cannot *prevent* attention from being captured by the singleton distractor, but rather it allows a fast and efficiently redirection of attention from the distractor to the target location.

The present findings are consistent with those in Kim and Cave 1999, which investigated the temporal interaction between top-down and bottom-up control of attention by means of probe RTs. Kim and Cave also used a task similar to that in Theeuwes 1992, where subjects searched for a shape singleton (a circle among diamonds) while an irrelevant color singleton distractor (a red element among green elements) was present. Either 60 or 150 msec after the presentation of the search display containing the target and singleton distractors, probes could appear at any of the locations. It was hypothesized that if the early preattentive processing is solely driven by bottom-up salience, as suggested by Theeuwes (1991, 1992), then the location of the salient singleton distractor should be attended first, and thus the probe RT at the distractor location should be faster than at any of the other locations in the short-SOA condition regardless of whether the unique feature is relevant. On the other hand, if top-down control is possible somewhat later in time, as the current experiments suggest, then in the late-SOA condition, attention should no longer be at the distractor location but instead at the location of the target singleton. For conditions in which target and distractor were locally unique (and therefore salient enough) Kim and Cave (1999) did indeed find these results. At an SOA of 60 msec, the probe RT at the location of the singleton distractor was about 20 msec faster than at the target singleton location. At an SOA of 150 msec, however, this pattern was reversed: the probe RT at the target location was about 15 msec faster than at the distractor location.

The current findings fit very well with those reported in Kim and Cave 1999, namely, that after 150 msec, attention is no longer at the location of the distractor but instead at the location of the target. In our experiment 1, we show that, at an SOA of at least 150 msec, the singleton distractor no longer interferes with search for the target singleten: there is no difference in RTs between the long-SOA conditions and the no-distractor condition. These findings both suggest that it takes somewhere between 100 and 150 msec to disengage attention from the location of the distractor and redirect it to the location of the target singleton.

The current results shed some new light on the findings obtained with the spatial cuing paradigm of Folk and colleagues (Folk, Remington, and Johnston 1992; Folk and Remington 1998) in which subjects have to ignore a cue that appears 150 msec before the search display. The critical finding is that a cue that does not match the top-down search goal (e.g., as in our experiments, the search goal is a shape singleton; the cue is a color singleton) does not affect RT, whereas a cue that does match the search goal slows search. The current findings and those in Kim and Cave 1999 show why with an SOA of 150 msec, a cue that does not match the search goal has no effect on RT: by the time the search display is presented, subjects are able to exert enough top-down control to allow a redirection of attention from the location of the distractor to the location of the target. The finding that there is an effect on RT in Folk and colleagues' experiments when the cue and target share the same defining property (e.g., the cue is red and the target is red) is not surprising because it is likely that disengagement and redirection of attention from the distractor location will take much longer when the distractor and target have the same defining property. It will be clear that this explanation of Folk and colleagues' data does not suggest anything like a "contingent capture" hypothesis, but merely confirms Theeuwes's stimulus-driven model of selection (1992), in which early processing is driven by bottom-up saliency factors. Note that our current findings and those of Kim and Cave (1999) also disconfirm a more recent notion put forward by Folk and Remington (1998), which suggests that irrelevant singletons do not capture spatial attention but merely cause a "filtering" cost. Both the effect of congruency of the letter inside the distractor, as found in our experiments, and spatial RT probe effect, as found in Kim and Cave 1999, clearly indicate that *spatial* attention was in fact captured by the location of the distractor.

Even though we suggested that the effect of the distractor at the later SOAs was reduced because of top-down control, the time course of the distractor effect could also be explained in a purely bottom-up fashion. Along these lines, it is assumed that attention is captured bottom-up by the most salient singleton and, after being disengaged from the most salient singleton, automatically reoriented to the next most salient singleton. If it takes about 150 msec to disengage and reorient attention, then it is not surprising that, at an SOA of 150 msec the interference effect was reduced. Note, however, that we assume that top-down control (i.e., knowing that one is looking for a diamond shape) does facilitate the disengagement of attention from the colored distractor singleton. After selecting the colored distractor, knowing that one is looking for a diamond and not for a red circle will most likely speed up the disengagement of attention and facilitate reorienting (see Theeuwes 1994b for a similar account).

We interpreted the current results in a strictly serial fashion, assuming that attention is first shifted to the most salient singleton and then to the next. Parallel processing models could also explain the current findings, assuming that on some trials the distractor finishes processing first, while on others, the target singleton finishes first. A purely parallel model, in which not only the two singletons are processed in parallel but all items are processed in parallel is somewhat less likely, given the findings of the control experiment (see note 1), which showed no congruency effect when a congruent or incongruent letter was placed in a nonsingleton. If all letters were processed in parallel, there should have been a clear congruency effect in the control study because the response-related letter inside a nonsingleton would have been processed at least as fast as any of the other letters in the display (possibly faster because of a top-down setting to look for this letter).

The current study indicates that during early preattentive processing, selection is driven bottom-up, that is, attention is captured by the most salient singleton present in the visual field. After attention is captured by the location of the singleton distractor, "attentive" processing exerts top-down activation that allows attention to be shifted elsewhere. The current model assumes that visual selection is the result of an interaction between goal-directed and stimulus-driven factors, consistent with models of visual search (e.g., Cave and Wolfe 1990; Wolfe 1994; Treisman and Sato 1990). Yet, unlike other models, the current model assumes that early preattentive parallel processing (assumed to calculate differences among stimulus features) is not accessible to top-down control. Only after an item has been selected does top-down processing help to speed up the disengagement of attention, allowing attention to be shifted to the next location.

NOTES

This research was supported by cooperative research agreement DAAAL01-96-2-0003 with the U.S. Army Research Laboratory. We thank Angela Glass and Meredith Minear for their assistance in conducting the studies.

1. To ensure that any congruency effect does indeed depend on attention being attracted to the colored singleton distractor, we ran a control study in which twelve subjects searched for a shape singleton while a congruent or incongruent letter was placed in one of the nonsingletons, instead of being placed in the colored distractor. There were no reliable effect of congruency on RT: $F(1,11) = 1.06$; $p = 0.32$; nor on error rate: $F(1,11) = 1.32$; $p = 0.27$.

REFERENCES

Bacon, W. F., and Egeth, H. E. (1994). Overriding stimulus-driven attention capture. *Perception and Psychophysics, 55,* 485–496.

Cave, K. R., and Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology, 22,* 225–271.

Caputo, G., and Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research, 38,* 669–689.

Egeth, H. E., Virzi, R. A., and Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32–39.

Egeth, H. E., and Yantis, S. (1997). Visual attention: control, representation and time course. *Annual Review of Psychology, 48,* 269–297.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter nonsearch task. *Perception and Psychophysics, 16,* 143–149

Eriksen, C. W., and Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics, 12,* 201–204.

Folk, C. M., and Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 847–858.

Folk, C. L., Remington, R. W., and Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1030–1044.

Joseph, J. S., and Optican, L. M. (1996). Involuntary attentional shifts due to orientation differences. *Perception and Psychophysics, 12,* 201–204.

Kaptein, N. A., Theeuwes, J., and van der Heijden, A. H. C. (1995). Search for a conjunctively defined target can be selectively limited to a color-defined subset of elements. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1053–1069.

Kawahara, J., and Toshima, T. (1996). Stimulus-driven control of attention: Evidence from visual search for moving target among static nontargets. *Japanese Journal of Psychonomic Science, 15,* 77–87.

Kim, M. S., and Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception and Psychophysics, 61,* 1009–1023.

Koch, C., and Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4,* 219–227.

Kumada, T. (1999). Limitations in attending to a feature value for overriding stimulus-driven interference. *Perception and Psychophysics, 61,* 61–79.

Maljkovic, V., and Nakayama, K. (1994). Priming of pop-out: 1. Role of features. *Memory and Cognition, 22,* 657–672.

Neil, W. T., and Valdes, L. A. (1996). Facilitatory and inhibitory aspects of attention. In A. F. Kramer, M. G. H. Coles and G. D. Logan (Eds.), *Converging operations in the study of visual attention,* pp. 77–106. Washington DC: American Psychological Association.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

Posner, M. I., and Cohen, Y. (1984). Components of visual orienting. In H. Bouma and D. Bouwhuis (Eds.), *Attention and Performance X,* pp. 531–556. Hillsdale, NJ: Erlbaum.

Sagi, D., and Julesz, B. (1985). Detection versus discrimination of visual orientation. *Perception, 14,* 619–628.

Theeuwes, J. (1991a). Cross-dimensional perceptual selectivity. *Perception and Psychophysics, 50,* 184–193.

Theeuwes, J. (1991b). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics, 49,* 83–90.

Theeuwes, J. (1991c). Categorization and identification of simultaneous targets. *Acta Psychologica, 76,* 73–86.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception and Psychophysics, 51,* 599–606.

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica, 83,* 93–154.

Theeuwes, J. (1994a). Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 799–806.

Theeuwes, J. (1994b). Endogenous and exogenous control of visual selection. *Perception, 23,* 429–440.

Theeuwes, J. (1996). Perceptual selectivity for color and form: On the nature of the interference effect. In A. F. Kramer, M. G. H. Coles and G. D. Logan (Eds.), *Converging operations in the study of visual attention,* pp. 297–314. Washington DC: American Psychological Association.

Theeuwes, J., and Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1342–1353.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., and Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1595–1608.

Todd, S., and Kramer, A. F. (1994). Attentional misguidance in visual search. *Perception and Psychophysics, 56,* 198–210.

Tipper, S. P., and Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitory effects of ignored primes. *Quarterly Journal of Experimental Psychology, 37A,* 591–611.

Treisman, A. M., and Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 451–478.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1,* 202–238.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2,* 156–161.

Yantis, S. (1996). Attentional capture in vision. In A. F. Kramer, M. G. H. Coles, and G. D. Logan (Eds.), *Converging operations in the study of visual attention,* pp. 45–76. Washington DC: American Psychological Association.

Yantis, S., and Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from selective search. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 601–621.